高校图书馆外文数据库资源建设的使用统计和分析思考*
——以复旦大学4个数据库的实践为例

■ 赵玮佳[1] 郝群[1] 张立彬[2]
[1]复旦大学文献信息中心 上海 200433 [2]南开大学图书馆 天津 300071

摘 要:[目的/意义]21世纪以来,数据库在高校图书馆采购资源中的比重不断上升,外文数据库是其中重要的电子资源来源。与此同时,面对各类监督审计工作的严格要求,工作人员在采购、建设外文数据库资源过程中也遇到诸多问题。以复旦大学图书馆为例,通过对已订购的Elsevier SD、Wiley、Springer和Taylor – ST等4个外文数据库的资源订购和使用情况的分析研究,为高校图书馆外文数据库建设决策提供思路和借鉴。[方法/过程]根据4个外文数据库的统计数据,分类归纳出馆藏资源概况、“双一流”学科建设的保障概况、年度利用情况和年度高被引无馆藏期刊情况。从这4个分析角度,基于扎根理论对数据进行编码,结合具体数据得到外文数据库综合使用情况表,并利用灰色关联分析方法对各指标的影响程度进行计算评估。[结果/结论]通过计算识别出影响复旦大学图书馆外文数据库使用最为关键的年度下载量、引用量、发文量、馆藏期刊总量和核心期刊占比5个指标,从学术成果补充计量、文献计量和资源保障3个角度为我国高校图书馆外文数据库建设提出优化建议,为后续其他高校信息资源建设研究提供方法的借鉴和重点把控的目标、依据和思路。

关键词:高校图书馆 外文数据库 资源建设 扎根理论 灰色关联分析
分类号:G250
DOI:10.13266/j.issn.0252 – 3116.2022.04.008

1 引言

随着数字图书馆、智慧图书馆的建设,图书馆用户使用电子资源的需求和频率大大增加,图书馆文献资源建设的重心也逐步从纸本转向电子资源。高校图书馆作为代表学校采购数据库的主体,购置电子资源数据库的费用也在每年的经费支出中占据“大部头”。根据教育部高等学校图书情报工作指导委员会发布的《2019年1165所高校图书馆电子资源购置费统计表》[1]、《2019年1263所高校图书馆文献资源购置费统计表》[2]和《2019年1278所高校图书馆年度总经费统计表》[3],可以看到,电子资源购置费排名前3位的高校图书馆——清华大学图书馆、北京大学图书馆和中山大学图书馆的电子资源购置费用分别为5 203.854 8万元、5 180.994 4万元和5 053.616 4万元,分别占文献资源购置费用的74.79%、69.63%和38.27%,占图书馆年度总经费的74.59%、54.70%和35.40%。香港中文大学图书馆在2019 – 2020年报[4]中也提到电子资源费用占物料支出的92%。我国在2022年QS世界大学综合排名前50位的8所综合类高校图书馆的外文数据库的数量和中文数据库数量的情况见表1。

从表1可知,这8所高等学府中,外文数据库的购买个数均占据到总数的一半以上。考虑到这8所综合类高校在中国学术研究领域的重要影响和示范作用,可知分析外文数据库的使用情况并对已购买的外文数据库的使用效果进行评估,进而对图书馆未来的外文数据库资源建设提供建议是很有必要的。鉴于此,本文结合已有研究和我国高校图书馆外文数据库的使用实际情况(以复旦大学图书馆为例),通过统计目前外文数据库的资源属性数据和利用数据,基于扎根理论对数据进行编码,并利用灰色关联分析方法,识别出关

* 本文系2016年度国家社会科学基金重点项目“网络时代开放教育资源引进与利用中的知识产权问题研究”(项目编号:16ATQ002)和2018年度教育部国内外高等教育比较研究项目“新时代国外优秀本科数学教材评价研究”研究成果之一。
作者简介:赵玮佳,硕士研究生;郝群,副研究馆员,通信作者,E-mail:libwjzx@fudan.edu.cn;张立彬,研究馆员。
收稿日期:2021 – 08 – 12 修回日期:2021 – 11 – 22 本文起止页码:73 – 86 本文责任编辑:王传清

表 1 国内部分高校图书馆已购电子数据库概况[5-12]

高校名称	中文电子数据库数/个	中文电子数据库占比/%	外文电子数据库数/个	外文电子数据库占比/%	总数/个
清华大学	99	40.41	146	59.59	245
北京大学	148	16.65	741	83.35	889
香港大学	162	18.66	706	81.34	868
复旦大学	114	35.19	210	64.81	324
香港科技大学	39	11.68	295	88.32	334
香港中文大学	219	18.14	988	81.86	1 207
浙江大学	76	25.94	217	74.06	293
上海交通大学	71	28.29	180	71.71	251

键指标，并据此提出高校图书馆外文数据库资源建设维度方面的优化策略和建议。

2 文献综述

2.1 高校图书馆外文数据库资源建设面临问题及解决方案理论综述

作为我国学习了解国外最新技术和科技动态的信息库，目前图书馆购买外文数据库主要有整库购买、分专辑或分学科购买两种方式。而现有外文数据库建设主要面临经费压力、资源结构问题、资源建设主动权问题、资源所有权问题[13]以及如何评价外文数据库资源建设状况问题。针对经费压力问题，王春生[14]认为主要是因为国内高校图书馆对数据商的销售策略和定价策略认识和理解不全面造成，因此图书馆应增强数据库采购人员的谈判能力，获取有利的采购条件。针对数据库资源结构不均衡以及因集团联盟采购而牺牲的合同期内调整馆藏资源的主动权问题，赵艳枝[15]通过学习国外图书馆成功抵制大宗交易的实践建议我国高校图书馆可组成谈判项目组，在重新谈判大宗交易条件的同时，考虑改变资源采购模式，如采用“按篇付费＋馆际互借＋开放存取”的模式采购电子期刊数据库。针对数据库资源所有权问题，首先在图书馆与数据商的“较量”中，图书馆应当以研究者的视角，通过对高校图书馆在数字资源引进中使用许可协议模型的调研与分析，使高校图书馆规避潜在的知识产权纠纷风险，提升数字资源采购的性价比[17]。另一方面，高校图书馆也可以打破现有格局，通过开放获取等方式，创造“学者创造学术成果→存储于机构知识库或网络学科平台→学者免费利用并创造新的学术成果”一种全新的学术交流模式[17]，从而减弱数据商对整个学术交流过程的控制。而在评价外文数据库建设方面，目前研究方法主要包括数据库使用绩效评价[18]、从用户数据维度对数字资源进行评价[19]、数据库的使用统计报告分析[20]及高校图书馆学术服务能力[21]和成效评估体系[22]的构建研究。

2.2 高校图书馆外文数据库资源建设实践综述

除了国内外学者在高校图书馆外文数据库建设理论方面开展研究，国内外高校也纷纷将理论运用于实践。对于如何提升外文资源检索体验问题，上海交通大学采取持续加强外文全文电子期刊模型应用的措施，同时完善资源地图，加大期刊及发文和引用资源整合，并引入大数据分析方法深入发掘学术文献使用规律，评估学术图书及学术出版机构影响力，研究外文学术文献引证工作，促进资源建设转型[24]。针对图书馆如何促进外文数据库馆藏资源的创新探索，阳昕等[24]建议高校图书馆在资源建设时可参考学科、院系在电子资源利用方面的关联特征和集群特征来多维度优化电子资源配置与采购。美国加州大学所有 10 个校区的图书馆已于 2021 年 7 月 27 日推出统一的 UC Library Search 检索和借阅系统[25]，该检索系统通过改进的图书馆间资源共享工具和简化的用户界面使 10 所图书馆内的资源（包括母语资源和外文资源）更容易被查找和借阅，从而最大限度地提高加州大学馆藏的利用率。代尔夫特理工大学和荷兰国家图书馆启动未来图书馆实验室，研究调查用户行为方式与活动，帮助他们更新颖和更有趣味地发现图书馆内多种语言的数字馆藏资源[26]。而在高校图书馆推动数据资源开放获取方面，更是有众多高校采取了行动，如牛津大学图书馆在 2022－2024 年规划中提到“增加对馆藏的访问并促进更大的发现和参与”，并将重点采取延长图书馆开放时长、扩展可借阅多语言电子化藏品的范围广度和时间上的跨度、加强社区的参与及图书馆与博物馆、档案馆和美术馆的合作等措施[27]。

综上可知，部分学者对高校外文电子期刊数据库资源采购的问题和策略进行了分析，还有部分学者基于特定视角或理论对高校外文数据库资源建设进行了

分析和评估,且已有部分高校将这些理论运用于实践。这些都为本研究的开展提供了良好的基础。但现有研究也存在一些不足:首先,尽管已有研究表明,从期刊[28-29]和机构[30-32]层面,研究机构的资源使用量与科研产出具有相关性,但对于中国高校而言,外语电子期刊因为其本身与中国学者母语的差异,导致对外文数据库的分析时不能不考虑其特殊性。例如,用户面对非母语文献时,更习惯将其先下载到本地再用翻译软件去阅读;或者一些外文电子期刊数据库在下载时设置推荐功能,向用户推荐系统判定的"相似"文章,而用户出于对非母语的不熟悉加之考虑到对文献的广度的需求,往往会接受系统的自动推荐,从而连带下载这些相似文章,然而这些文章下载后是否真的在用户科研过程中起到了作用,则无法判定。对于这些情况,其中的过程并不是完全透明的,而是更类似一个"灰色系统"——用户使用该外文数据库的研究过程并不完全明了,各外文数据库中期刊的浏览量、下载量等因素与对该用户真正起作用的文献类别、数量等因素的关系也并不完全清楚。因此,外文数据库的浏览与下载量并不足以完全表明它的使用效果。其次,目前高校图书馆在外文电子期刊数据库采选时所选用的策略往往是用单篇下载成本去定量评估,但正如上文所述,该视角过于单一,无法全面衡量用户对外文数据库的使用情况。为填补现有研究不足,本文将首先分析4个外文数据库资源属性数据,从而对图书馆外文电子期刊数据库的资源概况有一个宏观的了解,并基于这4个数据库的利用数据,利用灰色关联分析法从数据库统计的传统计量视角和统计视角中识别出外文数据库在资源建设时应考虑的关键指标,并结合计算结果给出本研究对于高校图书馆在外文数据库资源建设方面的思考。

3 复旦大学图书馆外文电子期刊库的统计分析与评估实践

3.1 实践数据

本次实践所用数据均来自复旦大学图书馆所订购的4个外文电子期刊数据库(Wiley、Taylor-ST、Elsevier SD和Springer)。需说明的是,本文选取4个外文数据库对其用户使用数据进行统计分析,虽然选取的数据库数量相比于市面上出现的外文数据库数量而言稍显薄弱,但这4个外文数据库收录环境与农业科学、化学、工程、计算及工程技术、物理学、数学、生命科学、医学、人文社科等多种学科的期刊、图书、工具书等,资源种类和数量庞大[33]。且该4个外文数据库商均是世界上公认的高品位学术出版公司,所收录的多是相关学科的核心资料,历史悠久,出版经验丰富,因此选择这4个外文数据库作为样本进行分析能够在一定程度上反映高校外文数据库的利用行为特征。统计数据维度包括该4个数据库2019年和2020年的使用量、拒绝访问量、2020-2021年购买的4个外文电子期刊库所收录的期刊数据、2018年至今复旦大学学者在Web of Science(WoS)的发文数据和2010年至今WoS中的引文数据。

(1)4个外文数据库使用量数据。以期刊为单位,字段包括期刊全称、数据库商、出版平台、DOI、ISSN、用户总下载量、下载用户ID数、年度总下载量、年度各月份下载量。

(2)4个外文数据库拒绝访问量数据。以期刊为单位,字段包括期刊全称、数据库商、学科分类、年度总拒访量、年度各月份拒访量。

(3)4个外文数据库收录期刊数据。以期刊为单位,字段包括期刊编号、ISSN、期刊全称、数据库商、纸本价(美元)、学科分类、是否有在线版、电子刊价格(美元)、2020年电子刊人民币价格、被引打分。

(4)4个外文数据库人民币价格(含捆绑刊)。以数据库为单位,字段包括数据库、名称、2019年定价、2020年定价。

(5)复旦大学学者发文和引文数据。以期刊文章为单位,原始数据为复旦大学学者2018-2020年发表的期刊文章,字段包括文章入藏号、引用的参考文献(其中经过数据整理,将参考文献逐篇拆分,可以用入藏号字段匹配回原文章,如果原文有30篇参考文献,则在参考文献拆分中被分为30行,受Excel行数限制,该列表放入了两个分表,分别为参考文献拆分1和参考文献拆分2,两者合并为所有的参考文献,共计130万余篇。参考文献中的期刊名多为缩写,期刊名繁简映射表可供参考转换)、引用的参考文献数、作者、语种、文献类型、摘要、WoS核心合集的被引频次计数、被引频次合计(WoS核心、Biosis Citation Index(BCI)和中国科学引文数据库(Chinese Science Citation Database,CSCD))、使用次数(最近180天)、使用次数(2013年至今)、出版商、ISSN、EISSN、出版日期、卷、期、DOI、页数、WoS类别、研究方向、公开访问指标符等。

3.2 4个外文数据库馆藏资源分析

3.2.1 4个外文数据库馆藏学科分布

利用Tableau软件,分析4个外文数据库的馆藏数

量对比、馆藏学科数量对比以及每个数据库中各学科馆藏的占比情况，并将结果以可视化方式展现，见图 1。

从图 1 中左侧可看出 4 个外文数据库馆藏量的区别，右侧可看出 4 个外文数据库的不同学科在复旦大学图书馆的馆藏情况。总体来看，馆藏中理学、医学和工学资源数量最为丰富。由图 1 中间的流向图可看出，在 Elsevier SD 数据库中，馆藏数量较多的是理学、工学、医学，而在哲学、教育学、文学、历史学和军事学领域并无馆藏；Springer 数据库的馆藏量是这 4 个外文数据库中最多的，而其中理学这一学科馆藏量与 Springer 库中的其他学科相比最为突出，其次是医学、工学、法学，而文科类相关学科，如哲学、历史学、文学等则没有相关馆藏；Taylor-ST 的馆藏总量是 4 个外文数据库中最少的，其中馆藏总量也主要分布在理学、工学、医学这 3 个学科领域；在 Wiley 数据库中，除了军事学和管理学类文献，其他学科领域的文献复旦大学图书馆均有涉及，其中理学、医学、教育学馆藏数量最多。

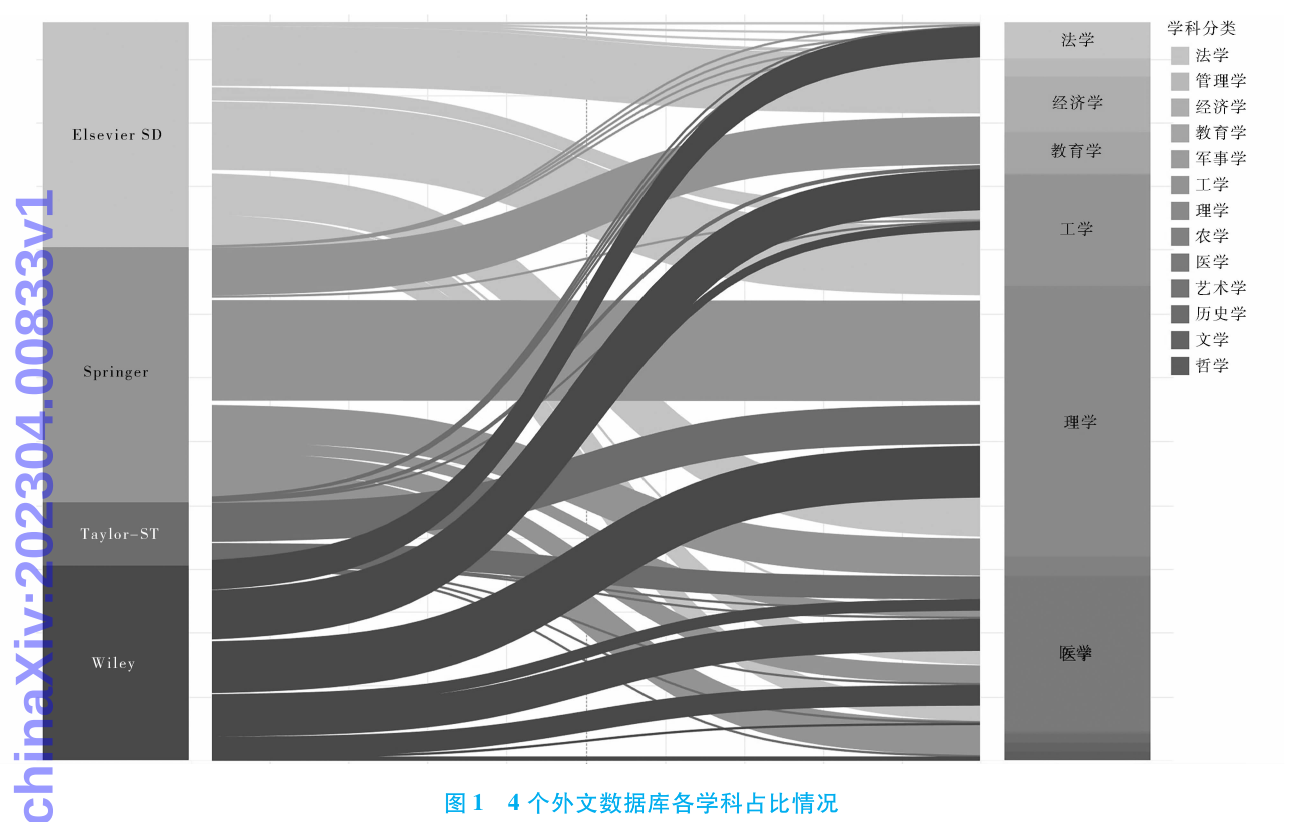


图 1　4 个外文数据库各学科占比情况

3.2.2　4 个外文数据库馆藏资源质量

通过分析图书馆 2020 - 2021 年购买的 4 个外文电子期刊库所收录的期刊数据，可以统计出各数据库馆藏外文电子期刊总量、各数据库间的重复刊数和净刊数量占比，并将 4 个外文电子期刊库的馆藏列表与核心期刊目录（JCR）进行比对[34]，了解 4 个数据库馆藏资源质量概况，具体如表 2 所示：

表 2　4 个数据库馆藏资源质量概况

数据库	馆藏期刊总量/种	重复刊数 /种	净刊占比/%	核心期刊数(JCR)/种	核心期刊(JCR)占比/%
Elsevier SD	1 677	0	100	1 216	72.51
Wiley	1 386	0	100	880	63.49
Springer	1 990	0	100	1 075	54.02
Taylor-ST	502	0	100	270	53.78

如表 2 所示，4 个外文电子期刊库虽然在购买期刊种数上有差异，但各个库之间都没有重复购刊，说明选库策略较为合理。

3.2.3　4 个外文数据库对双一流学科建设的保障概况

根据中华人民共和国教育部颁布的《“双一流”建

设学科名单》，复旦大学共有哲学、政治学、中国语言文学、中国史、数学、物理学、化学、生物学、生态学、材料科学与工程、环境科学与工程、基础医学、临床医学、中西医结合、药学、机械及航空航天和制造工程、现代语言学等共17个学科入选。该17个“双一流”学科所对应的学科门类见表3。

图书馆可以根据教育部学科门类，结合复旦大学“双一流”学科建设名单，计算4个库馆藏资源分别对应各学科门类的匹配期刊数、期刊匹配率、拒绝访问率、核心匹配数、核心保障率，从而可了解图书馆馆藏4个外文数据库对“双一流”学科建设的保障概况。

表3　4个外文电子期刊库对“双一流”学科建设的保障概况

复旦大学“双一流”学科	所属学科门类	指标	数据库名称			
			Elsevier SD	Wiley	Springer	Taylor-ST
哲学	哲学	期刊匹配数/种	0	72	0	0
		期刊匹配率/%	0	5.19	0	0
		核心匹配数/种	0	2	0	0
		核心保障率/%	0	0.14	0	0
政治学	法学	期刊匹配数/种	0	17	266	0
		期刊匹配率/%	0	1.23	13.37	0
		核心匹配数/种	0	0	4	0
		核心保障率/%	0	0	0.20	0
中国语言文学、现代语言学	文学	期刊匹配数/种	0	72	0	0
		期刊匹配率/%	0	5.19	0	0
		核心匹配数/种	0	2	0	0
		核心保障率/%	0	0.14	0	0
中国史	历史学	期刊匹配数/种	0	72	0	0
		期刊匹配率/%	0	5.19	0	0
		核心匹配数/种	0	2	0	0
		核心保障率/%	0	0.14	0	0
数学、物理学、化学、生物学、生态学	理学	期刊匹配数/种	562	422	831	312
		期刊匹配率/%	33.51	30.45	41.76	62.15
		核心匹配数/种	147	109	153	40
		核心保障率/%	8.77	7.86	7.69	7.97
材料科学与工程、环境科学与工程、机械及航空航天和制造工程	工学	期刊匹配数/种	332	76	296	176
		期刊匹配率/%	19.80	5.48	14.87	35.06
		核心匹配数/种	19	3	8	0
		核心保障率/%	1.13	0.22	0.40	0
基础医学、临床医学、中西医结合、药学	医学	期刊匹配数/种	491	332	386	14
		期刊匹配率/%	29.28	23.95	19.40	2.79
		核心匹配数/种	35	13	8	1
		核心保障率/%	2.09	0.94	0.40	0.20

如表3所示，从外文数据库角度纵向来看，复旦大学图书馆馆藏4个外文电子期刊数据库的学科各有侧重，Wiley在哲学、文学和历史学方面提供了重要文献资源，而在法学、工学两大学科门类中，Springer更胜一筹。Elsevier SD在理学、工学和医学方面核心期刊保障率均为最高，相比之下，Taylor-ST除了理学，在其余各学科文献资源支持方面较为逊色。从各个学科的资源保障率横向来看，馆藏4个外文数据库中理学、工学和医学3类学科的总体馆藏期刊数量较多，且理学的核心期刊保障率最高，其次是医学和工学；哲学、法学、文学和历史学4类学科则在这4个库中的馆藏期刊数较少，核心期刊保障率也相应偏低，说明目前图书馆在这4个外文数据库资源建设重点在理工医类。

3.3　用户利用数据分析

3.3.1　4个外文数据库年度利用情况

考虑到读者对外文数据库的使用行为是建立在自身需求和动机上的，因而读者使用外文数据库时所产生的使用数据可看作读者使用活动的集中体现。通过

分析用户使用数据及相互间进行交叉匹配，图书馆可以从不同维度对4个外文数据库资源列表、WoS发文统计、WoS引文统计、拒绝访问量等进行交叉匹配。图2展示了复旦大学研究人员2020年度在4个外文数据库所发文献类型的数量情况。从图2中可知，馆藏4个数据库中，复旦大学学者在Elsevier SD数据库中的发文量最多，其次是Wiley和Springer，发表在Taylor-ST数据库中的文献数量最少；且在这4个数据库中发文类型多为文章（Artical），综述（Review）和会议文摘（MeetingAbstract）类型的文献紧随其后。

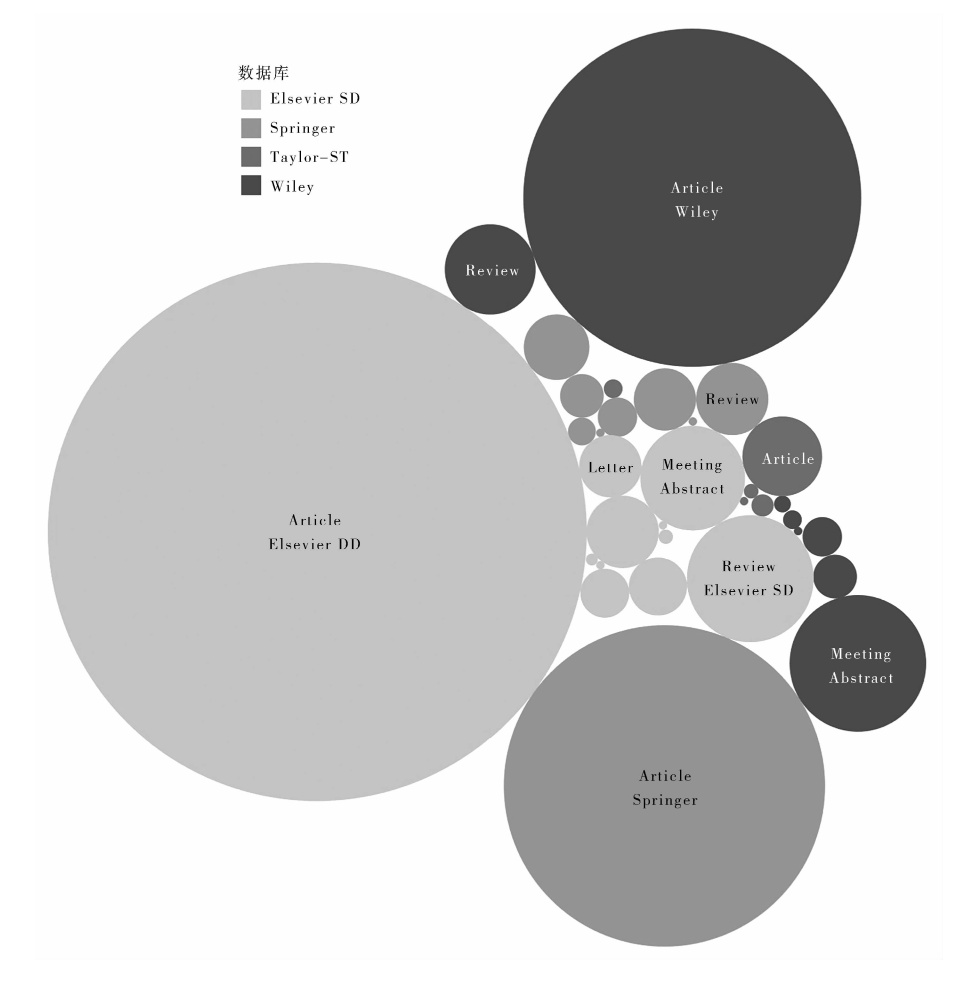


图2　复旦大学学者4个外文数据库2020年度发文类型

计算各数据库的年度下载量、单篇下载成本、0下载刊量占比、引用量、单篇引用成本、下载引用比以查看数据库的年度利用情况，如表4所示：

表4　复旦大学2020年度4个外文数据库已购期刊利用情况

数据库	下载量/次	发文量/篇	拒访量/次	数据库价格/元	单篇下载成本/元	0下载刊量/种	0下载刊量占比/%	引用量/次	单篇引用成本/元	下载引用比
ElsevierSD	2 645 774	4 700	2 044	4 187 545	1.58	75	4.47	2 416	1 733.26	1 095.11
Wiley	1 486 685	2 054	350	3 244 278	2.18	190	13.71	1 343	2 415.69	1 106.99
Springer	581 885	1 712	604	918 150	1.58	117	5.88	749	1 225.83	776.88
Taylor-ST	193 185	106	1 603	1 417 440	7.34	127	25.30	71	19 963.94	2 720.92

根据表4可知，在不同数据库中所需的单篇下载成本和引用成本，例如，对于Taylor-ST数据库而言，平均每引用1篇文章的成本约为1.9万元；而下载引用比则较为直接地反映出了读者从“利用信息”到“创造信息”的转化率，如在Springer数据库中，用户每下载约777篇文章，就有一篇会转化为发文时被引用的文章，转化效率为4个数据库之首。

3.3.2　年度有引用但无馆藏期刊

由于有引用无馆藏期刊所在的数据库也是订购策略调整时的重要依据[35]，因此图书馆可以以年为单

位,对该年度有引用却无馆藏的期刊,以及复旦大学学者有发文却无馆藏的期刊所在的数据库进行统计,来辅助订购决策的制定。将复旦大学图书馆2020年度有引用无馆藏的期刊按照被引用次数从高到低排序,被引次数最多的前三名期刊分别是 *JAMA-JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION*、*JOURNAL OF MATERIALS CHEMISTRY A* 和 *BMJ-BRITISH MEDICAL JOURNAL*,其所在的数据库如表5所示:

表5　复旦大学2020年度高被引无馆藏期刊(部分)

年度被引次数前3位无馆藏期刊	年度被引次数/次	所在数据库/出版商
JAMA-JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION	250	AMER MEDICAL ASSOC
JOURNAL OF MATERIALS CHEMISTRY A	53	ROYAL SOC CHEMISTRY
BMJ-BRITISH MEDICAL JOURNAL	49	BMJ PUBLISHING GROUP

3.4　使用情况表指标构建

本文通过研究高校4个外文数据库的使用统计数据,来构建读者对外文数据库综合使用情况表的指标维度,因此选取扎根理论法这一质性研究方法来进行这项探索性研究[36]。扎根理论是由格拉斯(B. Glaser)和斯特劳斯(A. Strauss)在1967年共同提出,该理论为研究人员提供了一套完整的从原始材料中归纳、构建理论的方法和步骤[37]。本文从收集整理复旦大学图书馆4个外文数据库使用统计数据开始,这些数据对外文数据库资源概况和读者利用行为进行了揭示,然后通过数据的性质进行编码区分、归类和综合来获得概念类属的内涵及之间的关联。

3.4.1　编码方法

由于本研究以诠释数据本身为主要目标,因此本文采取施特劳斯3种编码方式同时进行的策略,即开放编码、关联编码和选择编码。

(1)开放编码。在开放编码中,要求研究者凭借理论敏感性开放地挑选资料,在对数据处理的过程中忠于原始数据,构建简短而精确的初步代码体系。外文数据库使用统计中使用情况是通过许多事件来反映的,因此,本文采取逐个事件的方式进行初始编码,共获得概念类属13个,并以此作为关联编码的基础。开放编码过程如表6所示:

表6　开放编码过程

范畴	原始数据(部分)
馆藏期刊总量/种	数据库 Wiley 在本馆馆藏期刊有 *Journal of pharmacy practice and research*、*International journal of laboratory hematology*……
	数据库 Taylor-ST 在本馆馆藏期刊有 *Journal of liquid chromatography & related technologies*, *journal of business & economic statistics*……
净刊占比/%	数据库 Wiley 在本馆馆藏期刊有 *Journal of pharmacy practice and research*、*International journal of laboratory hematology*……
	数据库 Springer 在本馆馆藏期刊有 *Journal of materials science*、*Journal of geodesy*……
核心期刊占比/%	数据库 Wiley 在本馆馆藏期刊有 *Skin research and technology*、*International journal of laboratory hematology*……
	数据库 Taylor-ST 在本馆馆藏期刊有 *International reviews in physical chemistry*、*Carbon management*……
	数据库 Springer 在本馆馆藏期刊有 *Journal of materials science*、*Journal of geodesy*……
	数据库 Elsevier SD 在本馆馆藏期刊有 *International journal of antimicrobial agents*、*Journal of trace elements in medicine and biology*……
	核心期刊(JCR)有 *Skin research and technology*、*International reviews in physical chemistry*、*Journal of materials science*、*Journal of trace elements in medicine and biology*……
"双一流"学科门类期刊匹配率/%	复旦大学"双一流"学科门类有哲学、法学、理学……
	数据库 Elsevier SD 本馆馆藏期刊在理学有期刊 *Journal of fuel chemistry and technology*……
	数据库 Springer 本馆馆藏期刊在法学有期刊 *European business organization law review*……
"双一流"学科门类核心期刊保障率/%	复旦大学"双一流"学科门类有哲学、法学、理学……
	核心期刊(JCR)理学有 *Journal of energy chemistry*、*Journal of energy chemistry*……
	核心期刊(JCR)法学有 *Hague journal on the rule of law*、*European journal of psychology applied to legal context*……
	数据库 Elsevier SD 本馆馆藏期刊在理学有期刊 *Journal of energy chemistry*、*education for chemical engineers*……
	数据库 Springer 本馆馆藏期刊在法学有期刊 *Hague journal on the rule of law*……
年度下载量/次	数据库 Springer 馆藏期刊 *Journal of solid state electrochemistry* 在2020年1月下载量为35……
	数据库 Springer 馆藏期刊 *Journal of solid state electrochemistry* 在2020年2月下载量为30……
	数据库 Elsevier SD 馆藏期刊 *Academic pediatrics* 在2020年1月下载量为20……
	数据库 Elsevier SD 馆藏期刊 *Academic pediatrics* 在2020年2月下载量为12……

（续表 6）

范畴	原始数据(部分)
年度拒访量/次	数据库 Taylor-ST 馆藏期刊 *Journal of environmental planning and management* 在 2020 年拒访量为……
	数据库 Springer 馆藏期刊 *Journal of environmental planning and management* 在 2020 年拒访量为……
	数据库 Wiley 馆藏期刊 *Journal of interactive marketing* 在 2020 年拒访量为……
	数据库 ElsevierSD 馆藏期刊 *Journal of interactive marketing* 在 2020 年拒访量为……
单篇下载成本/元	数据库 ElsevierSD 馆藏期刊 *Academic pediatrics*2020 年下载量为 344 次……
	数据库 Elsevier SD 馆藏期刊 *Accident analysis &prevention*2020 年下载量为 939 次……
	数据库 Springer 馆藏期刊 *Journal of solid state electrochemistry*2020 年下载量为 688 次……
	数据库 Springer 馆藏期刊 *Journal of the European mathematical society*2020 年下载量为 1 次……
	数据库 ElsevierSD2020 年价格为 4 187 545 元……
	数据库 Springer2020 年价格为 918 150 元……
0 下载刊量占比/%	数据库 Springer 馆藏期刊 *Molecular medicine*2020 年下载量为 0 次……
	数据库 Springer 在本馆馆藏期刊有 *Molecular medicine*、*Environmental economics and policy studies*……
	数据库 ElsevierSD 馆藏期刊 *Journal of acupuncture and meridian studies*2020 年下载量为 0 次……
	数据库 ElsevierSD 在本馆馆藏期刊有 *HIV & AIDS review*、*Journal of acupuncture and meridian studies*……
下载引用比	2020 年 Chu AC 等人发表在 *Journal of population economics* 上的文章被复旦大学学者引用，该期刊已被 *Springer* 数据库馆藏收录……
	数据库 Springer 馆藏期刊 *Journal of solid state electrochemistry* 在 2020 年 1 月下载量为 35……
	数据库 Springer 馆藏期刊 *Journal of solid state electrochemistry* 在 2020 年 2 月下载量为 30……
引用量/次	2020 年 Li Z W 等人发表在 *Science of the total environment* 上的文章被复旦大学学者引用，该期刊已被 ElsevierSD 数据库馆藏收录……
	2020 年 Zhang B 等人发表在 *Journal of dispersion science and technology* 上的文章被复旦大学学者引用，该期刊已被 Taylor-ST 数据库馆藏收录……
单篇引用成本/元	2020 年 Li Z W 等人发表在 *Science of the total environment* 上的文章被复旦大学学者引用，该期刊已被 ElsevierSD 数据库馆藏收录……
	数据库 ElsevierSD2020 年价格为 4 187 545 元……
发文量/篇	2020 年 Cheng Z H 等人在 *Journal of liquid chromatography & related technologies* 上发表一篇文章，该期刊已被 Taylor-ST 数据库馆藏收录……
	2020 年 Zhang J M 等人在 *Journal of medical virology* 上发表一篇文章，该期刊已被 Wiley 数据库馆藏收录……

（2）关联编码。在开放编码的基础上，总结归纳形成概念类属并检视各类属之间的关系，进行关联编码。例如开放编码中“年度下载量”“年度拒访量”“单篇下载成本”“0 下载刊量占比”“下载引用比”均可归纳为“文献利用”等。最后形成 4 个关联编码概念。

（3）选择编码。通过系统处理关联式编码获得的概念类属之间的关系，进一步确定核心范畴和次要范畴，发现复旦大学图书馆 4 个外文数据库使用统计数据主要可以归纳为“资源保障”“学术成果补充计量”和“文献计量”3 个维度。资源保障包括外文数据库的“馆藏概况”“学科保障”；学术成果补充计量包括“文献利用”；文献计量包括“文献反馈”。本次研究三级编码结果见表 7。

（4）理论饱和度检验。本研究以预留的 2 个分析样本的有效数据进行理论饱和度检验，结果显示理论模型中的概念范畴已发展完备，除了分析复旦大学图书馆 4 个外文数据库使用统计的 3 个维度（资源保障、学术成果补充计量、文献计量）以外，没有再发现新的维度，且内部也未发现新的初始概念。因此，可以认为通过扎根理论得出的“复旦大学图书馆 4 个外文数据库使用统计分析维度”在理论上已经达到了饱和。

表 7　复旦大学 2020 年度外文数据库统计的编码表

开放编码	关联编码	选择编码
馆藏期刊总量/种	馆藏概况	资源保障
净刊占比/%		
核心期刊占比/%		
双一流学科门类期刊匹配率/%	学科保障	
双一流学科门类核心期刊保障率/%		
年度下载量/次	文献利用	学术成果补充计量
年度拒访量/次		
单篇下载成本/元		
0 下载刊量占比/%		
下载引用比		
引用量/次	文献反馈	文献计量
单篇引用成本/元		
发文量/篇		

3.4.2　关键指标识别

根据表 6 的开放编码，基于 4 个外文数据库利用数据，可以年为单位，得到复旦大学图书馆 4 个外文数据库的利用数据如表 8 所示：

表 8　4 个外文数据库利用数据统计

指标	Elsevier SD	Wiley	Springer	Taylor-ST
馆藏期刊总量/种	1 677	1 386	1 990	502
净刊占比/%	100	100	100	100
核心期刊占比/%	72.51	63.49	54.02	53.78
“双一流”学科门类期刊匹配率/%	0.825 9	0.663 1	0.894 0	1
“双一流”学科门类核心期刊保障率/%	0.119 9	0.091 6	0.086 9	0.081 7
年度下载量/次	2 645 774	1 486 685	581 885	193 185
年度拒访量/次	2 044	350	604	1603
单篇下载成本/元	1.58	2.18	1.58	7.34
0 下载刊量占比/%	4.47	13.71	5.88	25.3
下载引用比	1 095.11	1106.99	776.88	2720.92
引用量/次	2 416	1343	749	71
单篇引用成本/元	1 733.26	2 415.69	1 225.83	19 963.94
发文量/篇	4 700	2 054	1 712	106

由表 8 可知,各指标分别从不同角度反映了用户对图书馆订购的 4 个外文电子期刊数据库的使用情况。为了进一步明确表 8 中各指标的影响程度并识别出关键指标,本文采用灰色关联分析方法进行分析。

一般的抽象系统中都包含有多种因素,这些因素共同作用的结果决定了该系统的发展态势。但是人们往往希望在众多因素中识别出主要因素,知道哪些因素对系统发展影响大,哪些因素则对系统发展影响小;哪些因素起着推动系统发展作用并需要强化,而哪些因素又阻碍着系统发展需要加以抑制等。

数理统计中的回归分析、方差分析、主成分分析等都是用来进行系统分析的方法。但这些方法都有下述不足之处[38]:首先,这些分析方法都要求有大量数据做支撑,否则就难以找出统计规律;其次,这些方法均要求样本服从某个典型的概率分布,要求各因素数据与系统特征数据之间呈线性关系且各因素之间彼此无关,但这种要求在现实情况中往往难以满足;最后,使用这些分析方法可能出现量化结果与定性分析结果不符的现象,从而增加歪曲和颠倒系统的关系和规律的风险。尤其是当统计数据十分有限,而且现有数据灰度较大时,或由于人为原因,许多数据出现大起大落,没有典型的分布规律时,采用以上的数理统计方法往往难以奏效。

灰色关联分析方法则弥补了这样的缺憾。它对样本量的多少和样本有无规律并无明确要求,而且计算量小,十分方便,也不会出现量化结果与定性分析结果不符的情况[38]。灰色关联分析的基本思想是通过分析变量之间的几何形状的相似度进而判断其联系是否紧密的一种研究方法,该方法能够使各因素之间的“灰色”清晰化,进而能够确定事务的主要影响因素。该方法已在诸多领域得到了应用[39－41],具有可靠性和广泛适用性。本文采用灰色关联分析法,对表 8 进行计算,以明确表中各指标的影响程度并识别出关键指标,具体识别步骤如下:

(1)确定分析数列(母序列与子序列)。将表 8 中的各列最大值作为统计结果的最优值并作为母序列$\{X_0\}$,得到$\{X_0\}$ = {2 645 774,1 486 685,581 885,193 185}。子序列为馆藏期刊总量、净刊占比、核心期刊占比、“双一流”学科门类期刊匹配率、“双一流”学科门类核心期刊保障率、年度下载量、年度拒访量、单篇下载成本、0 下载刊量占比、下载引用比、引用量、单篇引用成本、发文量,分别记为 $X_1 - X_{13}$。

(2)对变量进行预处理,从而达到去量纲、缩小变量范围简化计算的目的:先求出每个指标的均值,再用该指标中的每个元素都除以其均值。结果如表 9 所示:

表 9　预处理后的矩阵

预处理	Elsevier SD	Wiley	Springer	Taylor-ST
X0	2.156 5	1.211 8	0.474 3	0.157 5
X1	1.207 6	0.998 0	1.432 9	0.361 5
X2	1.000 0	1.000 0	1.000 0	1.000 0
X3	1.189 7	1.041 7	0.886 3	0.882 4
X4	0.976 5	0.784 0	1.057 0	1.182 4
X5	1.261 3	0.964 3	0.914 9	0.859 5
X6	2.156 5	1.211 8	0.474 3	0.157 5
X7	1.777 0	0.304 3	0.525 1	1.393 6
X8	0.498 4	0.687 7	0.498 4	2.315 5
X9	0.362 2	1.111 0	0.476 5	2.050 2
X10	0.768 5	0.776 8	0.545 2	1.909 5
X11	2.110 5	1.173 2	0.654 3	0.062 0
X12	0.273 6	0.381 3	0.193 5	3.151 5
X13	2.193 2	0.958 5	0.798 9	0.049 5

(3)计算子序列中各个指标与母序列的关联系数。设 n 为样本数,m 为子序列个数,则 n 取 4,m 取 13。对每个指标分别计算$|X_{0k} - X_{ik}|$,$1 \leqslant i \leqslant 13, 1 \leqslant k \leqslant 4$,并记 $a = min(i)\ min(k)\ |X_{0k} - X_{ik}|$, $b = max(i)\ max(k)\ |X_{0k} - X_{ik}|$,计算后的矩阵见表 10。

(4)根据公式 $V(X_0(k), x_i(k)) = \frac{a + \rho b}{|X_0(k) - X_i(k)| + \rho b}$,$(i = 1,2,\cdots,m;\ k = 1,2,\cdots,n)$计算每个子序列和母序列的灰色关联度 V,其中 ρ 为分辨系数,一般为 0.5。求出的各指标的总体关联度见表 11。

表 10　$|X_{0k}-X_{ik}|$矩阵

	|X0 - X1|	|X0 - X2|	|X0 - X3|	|X0 - X4|	|X0 - X5|	|X0 - X6|	|X0 - X7|	|X0 - X8|	|X0 - X9|	|X0 - X10|	|X0 - X11|	|X0 - X12|	|X0 - X13|
	0.948 9	1.156 5	0.966 8	1.179 9	0.895 1	0	0.379 4	1.658 0	1.794 2	1.387 9	0.045 9	1.882 8	0.036 6
a = 0	0.213 7	0.211 7	0.170 0	0.427 7	0.247 4	0	0.907 4	0.524 0	0.100 7	0.434 9	0.038 5	0.830 4	0.253 2
	0.958 6	0.525 7	0.412 0	0.582 7	0.440 5	0	0.050 8	0.024 1	0.002 2	0.070 9	0.180 0	0.280 7	0.324 6
b = 2.994	0.204 0	0.842 5	0.724 9	1.024 9	0.702 0	0	1.236 1	2.157 9	1.892 7	1.751 9	0.095 4	2.994 0	0.107 9

表 11　灰色关联度计算结果

	γ(X0(k),X1(k))	γ(X0(k),X2(k))	γ(X0(k),X3(k))	γ(X0(k),X4(k))	灰色关联度 γ(X0,Xi)
X1	0.612 0	0.875 1	0.609 6	0.880 1	0.744 2
X2	0.564 2	0.876 1	0.740 1	0.639 9	0.705 1
X3	0.607 6	0.898 0	0.784 2	0.673 8	0.740 9
X4	0.559 2	0.777 8	0.719 8	0.593 6	0.662 6
X5	0.625 8	0.858 1	0.772 6	0.680 8	0.734 3
X6	1.000 0	1.000 0	1.000 0	1.000 0	1.000 0
X7	0.797 8	0.622 6	0.967 2	0.547 7	0.733 8
X8	0.474 5	0.740 7	0.984 1	0.409 6	0.652 2
X9	0.454 8	0.937 0	0.998 5	0.441 6	0.708 0
X10	0.518 9	0.774 9	0.954 8	0.460 8	0.677 3
X11	0.970 2	0.974 9	0.892 7	0.940 1	0.944 4
X12	0.442 9	0.643 2	0.842 1	0.333 3	0.565 4
X13	0.976 1	0.855 3	0.821 8	0.932 7	0.896 5

(5)关键指标识别。表 11 中的 γ(X0,Xi)列的数值越大,表示该指标越重要,据此根据 γ(X0,Xi)的大小进行排序即可确定关键评价指标。根据表 11 计算结果可知,对于复旦大学而言,影响 Elsevier SD、Wiley、Springer 和 Taylor-ST 这 4 个外文电子期刊库使用效果的前 5 个指标依次是年度下载量(1.000 0)、引用量(0.944 4)、发文量(0.896 5)、馆藏期刊总量(0.744 2)和核心期刊占比(0.740 9)。

3.5　结论与讨论

已有研究表明,灰色关联度值大于 0.5 表示比较数列对参考数据列的影响程度较大[42]。本文通过对构建的复旦大学用户外文电子期刊库使用情况表进行灰色关联分析,计算出了各指标的影响程度并识别出关键指标,从计算结果来看,表 11 中所有指标的总体关联度值均高于 0.5,说明指标选取得较为合理,并得出如下结论:

(1)在补充计量构面中,“年度下载量”是所有指标关联度值中最大的指标,说明图书馆在考虑外文数据库资源建设时,首要需要关注的关键指标是该数据库的年度下载量。学术论文是高校最直接最重要的科研产出,而数据库中的文献资源对于本校师生开展学术研究的支撑程度是影响高校购买的主要原因。科研人员通过浏览数据库后对里面的文章进行全文下载是开展后续研究的第一步,也是读者使用数据库最明显的行为。因此,对于学校已购买的外文数据库进行评价时可将该数据库年度下载量作为一个基础的客观评价指标,这也与以往图书馆数据库采选经验一致。除此之外,“年度拒访量”和“0 下载刊量”指标的灰色关联度值都超过了 0.7。由于科研过程和用户使用外文数据库行为的复杂性,不同学科背景的外文数据库的年度下载量也会出现截然不同的情况,即使是同一机构下的不同外文数据库的年度下载量之间也存在出入,因此,还需要在考虑外文数据库年度下载量的基础上结合其他重要指标。由于有拒访量说明读者有需求而图书馆尚未满足,0 下载刊量说明图书馆提供的资源服务读者可能并不需要,因此这二者也是订购策略调整时的重要依据。

(2)在文献计量构面中,“引用量”指标的灰色关联度值超过了 0.9,“发文量”指标的灰色关联度值也达到了 0.896 5,说明高校图书馆在建设外文数据库资源中,数据库中的文献资料被本校师生引用数量与本校师生所发表的文章被数据库收录数量这两个指标极为重要。数据库中的文献资料被用户下载后,若在后续科研过程中被利用,则最直接的表现方式就是在研究成果中对该文章进行了引用,这也表明该数据库的内容资料促进了本校科研人员的创新思维的产生。对

于资深学者而言，其发表的论文数量越多，他在本学科内发表高影响力论文的概率就越大[43]。由于学术论文是高校最主要的科研产出，本校师生在某一外文数据库中的发文量就会对本校科研人员的研究产生影响。因此，高校图书馆作为支撑师生科研的文献信息中心，在进行本馆的资源建设时，也必须考虑到该外文数据库对本校科研著作的收录情况。同时从计算结果中还能看到，虽然灰色关联度值排名第二高和第三高的两个指标(引用量和发文量)均属于文献计量构面，但"单篇引用成本"的灰色关联度值是最低的，因此在评估用户使用情况、制定外文数据库订购策略时，可优先考虑其他指标。

(3)在资源保障构面中，"馆藏期刊总量"指标与"核心期刊占比"指标的灰色关联度值均高于0.74，说明图书馆外文数据库在建设中仍然要考虑到馆藏总体量和质量。馆藏期刊总量代表本校使用的该外文数据库涵盖的内容范围的广度，而"核心期刊占比"则更进一步揭示了该外文数据库内容的质量深度，二者对于本校学者的科研支持缺一不可。对此，高校图书馆在进行馆藏资源建设布局时，除了以上提到的指标，也需要根据学校自身的学科建设规划考虑外文数据库内容的广度和深度。近年来随着国家"双一流"建设的提出与推进，"坚持以学科为基础"的建设原则亦成为高校图书馆资源建设的重要指导思想[44]。从计算结果中也看出，"双一流"学科门类核心期刊保障率和"双一流"学科门类期刊匹配率的灰色关联度值高达0.734 3和0.662 6，说明高校要建设一流的学科，就需要有与学校学科发展阶段相匹配的、优质的(核心)文献资源，这是基本条件与重要支撑。

4 对高校图书馆外文数据库建设的启示与思考

4.1 数据驱动分析馆藏资源

《中共中央关于制定国民经济和社会发展第十四个五年规划和二〇三五年远景目标的建议》[45]中提出"推动公共文化数字化建设"。高校图书馆是学校内的公共文献信息资源中心，因此加强信息资源建设是图书馆亘古不变的目标，这就要求图书馆要立足当前时代发展背景，考虑到不同资源特点和读者利用情况等因素构建信息资源体系。在馆藏资源的数据化转型过程中，对于外文数据库的采选和馆藏资源建设，图书馆应以服务师生为中心，以支持学校学科建设尤其是保障学校双一流学科建设为目标。因此，若能以数据为驱动对馆藏资源的建设现状和读者利用情况进行分析，就可以对馆藏资源分布、质量、对各学科的保障情况和利用情况有所了解；再以外文数据库的使用情况表为基础，计算寻找数据库资源建设过程中应重点把控的目标，就可以减少馆内外文数据库资源建设在结构设置、采选、评估等方面的阻碍，避免图书馆在外文数据库资源建设管理过程中由于外文资源语言差异而导致的"眉毛胡子一把抓"困境，也有助于加强馆藏资源管理过程中的理论依据和智慧程度。

4.2 挖掘外文数据库下载量、引用量、发文量及相关数据的研究价值

图书馆外文数据库留有用户利用过的"痕迹"，本文通过使用灰色关联分析发现，图书馆在建设外文数据库资源时，不仅需要注意所购买的馆藏期刊总量和馆藏核心期刊占比，更应增强对外文数据库年度下载量、引用量和发文量及其相关数据的分析与研究，发挥这些数据在资源建设中学术评价、信息检索、科学知识梳理及可视化、学科规划发展等方面的研究价值。从下载量及相关数据方面，外文学术文献下载行为是科研人员文献检索行为的重要一环，高校图书馆关注下载量对于加强本馆的外文数据库建设具有重要意义。首先，预测用户对馆藏外文数据库的文献下载量有助于深度理解用户在外文数据库的检索行为，为学校的学术资源检索平台优化检索结果、重构排序提供依据，从而提升检索系统的服务质量[46]。其次，探究下载量与其他因素关联性，如下载量与发文量或被引频次的关联性等有助于图书馆确定外文数据库绩效评价指标，提升图书馆决策的科学性。最后，关注下载量与用户偏好研究等除了可用于图书馆外文数据库用户使用情况的可视化呈现，也有助于图书馆对用户利用外文数据库的行为有更深度科学的理解，进而对用户进行更精准的画像并提供个性化服务。从引用量及相关数据方面，以外文数据库引用量为出发点，开展对被引文献的引用频次、引用位置、引用文本主题的挖掘研究，可据此对外文数据库进行描述性统计并助力引文网络的构建。从外文数据库的发文量及相关数据方面，图书馆可借助多种可视化工具多维度对某领域进行定量和研究主题聚类分析，揭示该领域发展态势、研究力量和研究热点，并从学科规划发展角度进行外文数据库建设。

4.3 数字资源共享共建

当前我国高校图书馆外文数据库建设面临着经费

压力、资源结构等问题[13]，而图书馆之间的资源共享与共建能够提高信息资源的整合力度与完善程度，这是高校图书馆在智能时代对外文数据库资源建设提出的进一步要求。高校图书馆在制定数据库资源建设策略时，除了关注自身已购买的外文数据库，还应考虑到与其他高校图书馆、公共图书馆及政府企业间的资源共享与共建，深化合作。首先，高校图书馆可通过参与高校图书馆联盟的方式，在各高校间建设统一的外文数据资源平台对数据进行共享及统一开发利用，为联盟内各高校图书馆数据资源的利用和分析提供有效保障；同时应考虑到不同学科背景下的读者使用资源类型、资源使用习惯的特点，在资源共建合作时有所侧重。其次，应积极探索具有本馆独特风格特色的文献资源。如利用本校重点学科的优势，对已有外文馆藏资源进行深入分析，并对这些资源进行收集整理，从而构建一套具有本馆特色的资源体系用于自用和共享。再次，高校图书馆还需注重共享中的隐私保护及共享资源的所有权保护，可采取加强图书馆隐私保护监管制度、建立科学数据共享中的隐私保护监管制度、建立隐私风险评估机制等隐私保护配套措施[47]，避免不必要的资源权益纠纷。最后，其他高校图书馆也可依据本文思路方法，从数据角度实时对本馆外文资源共建情况进行分析把控。

5 结语

本文依据用户信息行为理论，结合已有研究和高校馆藏资源建设的实际特点，通过整理归纳不同角度的外文数据库用户使用的具体数据，对复旦大学图书馆 2020 年度 4 个外文数据库用户使用情况进行分析，并依据扎根理论编码总结出外文数据库综合使用情况表，最后利用灰色关联分析方法对表中各指标的影响程度进行计算评估并识别出关键指标并给出高校图书馆外文数据库建设策略建议。本文的特色和贡献在于：①基于用户对外文数据库利用情况这一情境扩充了图书馆资源建设分析角度并以复旦大学图书馆为例进行实证，从而进一步确定了该方法的实用性和适用性。②利用扎根理论对用户利用外文数据库的统计数据进行编码区分、归类和综合，构成复旦大学用户外文数据库使用情况表，作为后续分析的依据。③利用灰色关联分析法成功识别关键指标，为后续高校信息资源建设研究提供了方法的借鉴和重点把控的目标、依据和思路。

研究也存在一些不足：首先，指标可以更细化，如学科分类可以细化至一级、二级学科；拒访量可以具体到学科分类下分别计算。其次，研究所使用的样本量较小，仅为 4 个外文电子期刊库，未来可进一步将更多数据库纳入调研范围（例如开放存取数据库等），扩大样本规模。

参考文献：

[1] 教育部高等学校图书情报工作指导委员会. 2019 年 1165 所高校图书馆电子资源购置费统计表[EB/OL]. [2021 -07 -30]. http://www.scal.edu.cn/tjpg/202011160231.
[2] 教育部高等学校图书情报工作指导委员会. 2019 年 1263 所高校图书馆文献资源购置费统计表[EB/OL]. [2021 -07 -30]. http://www.scal.edu.cn/tjpg/202011160231.
[3] 教育部高等学校图书情报工作指导委员会. 2019 年 1278 所高校图书馆年度总经费统计表[EB/OL]. [2021 -07 -30]. http://www.scal.edu.cn/tjpg/202011160231.
[4] 香港中文大学图书馆. 2019 -2020 年报[EB/OL]. [2021 -07 -30]. https://www.lib.cuhk.edu.hk/sites/cuhk/files/page/about/inside_the_library/Annual_Report_tc_19 -20.pdf.
[5] 清华大学图书馆. 清华大学数据库导航系统 -x server 版[EB/OL]. [2021 -07 -28]. http://nav.lib.tsinghua.edu.cn/xport/dbdh.htm.
[6] 北京大学图书馆. 数据库导航[EB/OL]. [2021 -07 -28]. https://dbnav.lib.pku.edu.cn/.
[7] 复旦大学图书馆. 学术资源门户[EB/OL]. [2021 -07 -28]. http://202.120.227.56:8331/V/24N8BV6BNHSBTTGCD4IDD74LELC4915ANHB42775HC9H4ALFX1-01000? FUNC = FIND-DB-1-LCL&pds_handle = GUEST.
[8] 上海交通大学图书馆. 数据库列表[EB/OL]. [2021 -07 -28]. http://www.lib.sjtu.edu.cn/f/database/database.shtml.
[9] 香港大学图书馆. Databases[EB/OL]. [2021 -10 -30]. https://libguides.lib.hku.hk/az.php.
[10] 香港科技大学图书馆. All databases - HKUST library[EB/OL]. [2021 -10 -30]. https://library.ust.hk/collections-resources/databases/all-databases/.
[11] 香港中文大学图书馆. A-Z Databases[EB/OL]. [2021 -10 -30]. https://libguides.lib.cuhk.edu.hk/az.php.
[12] 浙江大学图书馆. 数据库导航[EB/OL]. [2021 -10 -30]. http://210.32.137.90/s/lib/libtb/.
[13] 苏建华. 国内高校图书馆外文电子期刊采购的困境及应对策略[J]. 新世纪图书馆,2016(7): 48 -50.
[14] 王春生. “程焕文之问”:原因与出路——对外文期刊数据库采购工作的思考[J]. 图书馆论坛,2015, 35(5): 1 -5.
[15] 赵艳枝. 国外图书馆应对学术期刊大宗交易危机的策略及对我国的启示[J]. 图书馆工作与研究,2015(12): 61 -64.
[16] 方诗雨,郝群,张立彬. 美国高校图书馆数字资源引进中使用许

可协议模型的调研与分析[J]. 情报理论与实践,2021,44(8):187－195,158.

[17] 胡赛. 高校图书馆电子期刊交易困境及应对策略——基于数据商的视角[J]. 高校图书馆工作,2018, 38(5): 45－49.

[18] 向林芳,秦霞. 高校图书馆外文电子期刊全文数据库使用绩效评价——以华南理工大学图书馆为例[J]. 情报探索,2020(10): 95－101.

[19] 毕艳芳,李泰峰. 用户数据维度的数字资源评价方法及实例——以电子科技大学图书馆为例[J]. 图书情报工作,2017, 61(22): 82－88.

[20] 汤罡辉,程焕文. Wiley期刊库使用统计报告分析——以中山大学图书馆为例[J]. 图书情报工作,2007,51(4):136－138,149.

[21] 王琼,吴娱,吴英梅. 国外学术图书馆服务能力评价体系研究[J]. 大学图书馆学报,2009, 27(2): 85－90.

[22] 向远媛,温国强. 大学图书馆成效评估及其指标体系构建探索[J]. 大学图书馆学报,2011, 29(3): 11－16.

[23] 上海交通大学图书馆年度报告. 内容详情[EB/OL]. [2021－07－28]. http://www.lib.sjtu.edu.cn/f/content/detail.shtml;JSESSIONID = 30cb4967-37eb-40f9-b5e2-8b464f6b9d43? id = 6302&lang = zh-cn.

[24] 阳昕, 张敏, 廖剑岚,等. 社会网络视角下的高校图书馆电子资源利用研究——以复旦大学图书馆为例[J]. 图书情报工作, 2021, 65(15): 91－99.

[25] UC library search. UC libraries[EB/OL]. [2021－07－30]. https://libraries.universityofcalifornia.edu/uclibrarysearch.

[26] TU Delft and the National Library of the Netherlands launch future libraries lab[EB/OL]. [2021－07－30]. https://www.tudelft.nl/en/2020/tu-delft/tu-delft-and-the-national-library-of-the-netherlands-launch-future-libraries-lab.

[27] University of Oxford glam digital programme. Gardens, libraries &museums[EB/OL]. [2021－07－30]. https://www.glam.ox.ac.uk/digital-strategy.

[28] 王伟,王琳. 高校图书馆外文电子期刊利用行为与科研产出的相关性研究[J]. 情报探索,2016(12): 11－16.

[29] 刘佳音. 国内高校数字资源利用效能综合评价研究[D]. 武汉:武汉大学, 2013.

[30] BOUKACEM-ZEGHMOURI C, BADORP, LAFOUGE T, et al. Relationships between consumption, publication and impact in French universities in a value perspective: a bibliometric analysis[J]. Scientometrics,2016,106:263－280.

[31] NICHOLAS D, ROWLANDS I, HUNTINGTON P, et al. E-journals: their use, value and impact [EB/OL]. [2021－06－30]. https://www.researchgate.net/publication/266034523_E-jourals_their_use_value_and_impact.

[32] JUNG Y, KIM J, SO M, et al. Statistical relationships between journal use and research output at academic institutions in South Korea [J]. Scientometrics,2015,103:751－777.

[33] 刘佳音. 高校图书馆电子资源使用与用户检索行为统计分析——以ScienceDirect数据库为例[J]. 大学图书馆学报,2012,30(2):81－86.

[34] 庄纪林,陈凌,姚晓霞. 外文期刊数据库内容质量指标的量化研究[J]. 图书情报工作,2021, 65(2): 54－64.

[35] 李娜. 电子资源管理与评估——以上海财经大学图书馆的外文电子期刊为例[J]. 图书情报工作,2018, 62(15): 60－62.

[36] 米梓源,张立彬,丁可宁. 基于扎根理论的高校图书馆提升电子图书利用效率分析[J]. 情报科学, 2019, 37(7): 48－55.

[37] 柯平,张文亮,李西宁,等. 基于扎根理论的馆员对公共图书馆组织文化感知研究[J]. 中国图书馆学报,2014, 40(3): 37－49.

[38] 刘思峰,党耀国,等. 灰色系统理论及其应用(第5版)[M]. 北京:科学出版社,2010.

[39] 张新香,段燕红. 基于学习者在线评论文本的MOOC质量评判——以"中国大学MOOC"网的在线评论文本为例[J]. 现代教育技术,2020, 30(9): 56－63.

[40] 张坤,王雪,李力. 高校图书馆微信公众号服务质量评价指标体系的构建与分析[J/OL]. 图书馆建设:1－11[2021－08－11]. http://kns.cnki.net/kcms/detail/23.1331.G2.20210423.1455.008.html.

[41] 牛勇平. 关于公共图书馆效率影响因素的灰色关联分析[J]. 图书馆杂志,2012, 31(2): 5－10.

[42] 蒲筱哥,刘礼明. 数据库绩效的变异系数与灰色关联度分析及实证研究[J]. 图书情报工作,2014, 58(14): 71－78.

[43] LARIVIERE V, COSTAS R. How many is too many? On therelationship between research productivity andimpact[J]. PLoS One, 2016, 11(9): e0162709.

[44] 张宇娥,柯佳秀,杜丽. 高校学科期刊资源保障优先级方法与实证分析[J]. 图书情报工作,2020, 64(8): 71－77.

[45] 中共中央关于制定国民经济和社会发展第十四个五年规划和二〇三五年远景目标的建议[EB/OL]. [2021－01－25]. http://china.cnr.cn/news/20201104/t20201104_525318774.shtml.

[46] 谢豪,吴雪华,陈茜,等. 融合多维特征的学术文献下载行为预测研究[J]. 图书情报工作,2021,65(12):112－121.

[47] 朱贝,王传清. 我国科学数据共享中的隐私治理探析[J]. 图书情报工作, 2020, 64(22): 37－47.

作者贡献说明：

赵玮佳：论文撰写；

郝群：题目设计、写作指导、部分内容撰写；

张立彬：体系架构、部分内容撰写、论文修改。

图书情报工作
第 66 卷 第 4 期 2022 年 2 月

The Usage Statistics and Analytical Thinking of Foreign Language Database Resource Construction in University Libraries: Taking the Practice of Four Databases in Fudan University as an Example

Zhao Weijia[1] Hao Qun[1] Zhang Libin[2]
[1] Fudan University Library, Shanghai 200433
[2] Nankai University Library, Tianjin 300071

Abstract: [**Purpose/significance**] Since the 21st century, the proportion of databases in the procurement resources of university libraries has been increasing, and foreign language databases have been an important source of electronic resources. At the same time, facing the strict requirements of various supervision and auditing work, the staff also encountered many problems in the process of purchasing and constructing foreign language database resources. Taking Fudan University Library as an example, through the analysis and research on the resource ordering and usage condition of the four foreign language databases of Elsevier SD, Wiley, Springer and Taylor-ST ordered by Fudan University Library, it provides ideas and references for the decision-making of foreign language database construction in university libraries. [**Method/process**] Based on the statistical data of the four foreign language databases, this paper classified and summarized the overview of the collection resources, the guarantee overview of the double-class discipline construction, the annual utilization status and the annual high-cited non-collection journals. From these four analysis perspectives, the data was coded based on the grounded theory, and then combined with specific data to obtain a comprehensive usage table of foreign language databases, and the gray relational analysis method was used to calculate and evaluate the influence of each indicator. [**Result/conclusion**] Through calculations, this paper identifies the five most critical annual downloads, citations, publications, total number of periodicals in the collection, and the proportion of core periodicals that affect the use of the foreign language databases in Fudan University Library. From three perspectives of bibliometrics of academic achievements, literature measurement and resource guarantee, this paper provides optimization suggestions for the construction of foreign language databases in Chinese university libraries, and provides methodological reference and key targets, basis and ideas for subsequent research on the construction of information resources in other universities.

Keywords: university library foreign language database resource construction grounded theory grey relational analysis